Optimal Replication Based on Optimal Path Hops for Opportunistic Networks

Andre Ippisch, Salem Sati and Kalman Graffi
Technology of Social Networks, Heinrich Heine University of Düsseldorf
Universitätsstrasse 1, 40225 Düsseldorf, Germany
Email: {ippisch, sati, graffi}@cs.uni-duesseldorf.de

***Abstract*—Opportunistic Networks are highly mobile networks which may lack a reliable path between some source and destination. Therefore, this type of network uses *Store, Carry and Forward* and delivers messages based on hop-by-hop routing. Epidemic is the simplest routing protocol for Opportunistic Networks, as it replicates messages to all encountered nodes. For messages spread by Epidemic replication, we study and analyze the trade-off between the messages' delivery ratio, delay, and overhead. We consider that all members of the network, at any given time, know the amount of relay nodes that pass the message and each message carrier knows the amount of infected nodes that already have acquired the message. We address the problem of deriving the optimal closed-loop control for the replication strategy in our network. We draft this issue as a controlled, discrete-time and finite-state *Markov Chain* with an *All Hops Optimal Path* formulation. In real life scenarios, however, due to the intermittent correspondence in Opportunistic Networks the nodes do not have insight of the network's global state. We try to solve this issue by obtaining an Ordinary Differential Equation approximation of the *Markov Chain* for the replication process of messages. Furthermore, our model considers *All Hops Optimal Path* as network graph analysis for the optimally controlled replication. Lastly, we present the performance evaluation of these replication control conditions in finite networks. Our results show that this proposed *Optimal Replication Based on Optimal Path Hops* performs better than the omni-directional and contact-based Epidemic routing replication.**

Keywords: Opportunistic Networks, Optimal Path, Markov Chain, Epidemic Routing, Optimal Replication.

I. Introduction

Opportunistic Networks (OppNets) are Delay Tolerant Networks (DTNs) which lack end-to-end paths between nodes. OppNets can, among other forms, consist of human-carried devices such as smartphones and smart devices [1]. Due to the human aspect, such a network includes high mobility which leads to an unreliable communication between nodes. The DTNs' and OppNets' environment operations are decentralized and the nodes have strong resource constraints. Among other things, relevant ones are energy and storage, as the networks are reliant on *Store, Carry and Forward* routing. However the potential lack of sufficient resources, source relay nodes need to control the transfer of messages when searching for the destination. One way to do this is limiting the maximum number of hops in the path on which the message travels through the network. A routing scheme that would limit the message's path to a maximum number of hops, more precise the optimal path hop limit, is desired. But determining the optimal hops of the path is neither straightforward nor feasible. Where a large number of hops tends to increase replication and overhead, it also increases the chance of finding the message destination. Thereby decreasing replication by shortening the number of hops leads to delivery loss or at least the delivery of the message with a high delay. In recent studies [2] it has been recognized that both, mobility and the network topology graph, can be exploited to improve routing performance. Specifically, in OppNet environments, to overcome mobility issues, the messages are routed from one hop to another based on additional information of the path. An originator distributes its message by replicating it to relay nodes for additional routes or paths between source and destination. Most message replication algorithms for OppNets commonly save multiple copies of the message at different nodes. These message copies attempt to increase the message delivery ratio, but this added replication leads to significant energy consumption, which shortens the life span of the battery-powered devices forming the OppNet. The simplest routing protocol is Epidemic [3] which floods the messages to all encountered nodes. Therefore, for the delivery of messages, Epidemic takes up more of the system's bandwidth, as well as the buffer space and energy of each node than necessary. Recently, many pieces of research [2] are devoted to opportunistic replication schemes for reducing the copies of each message in the network while staying relatively close to the desired delivery ratio.

In this work, we introduce *Optimal Replication Based on Optimal Path Hops (ORBOPH)*, a model which integrates an optimal delivery probability based on *Markov Chain* modeling combined with optimal path hops based on the *All Hops Optimal Path (AHOP)* formulation [4]. Our model aims for specifying the message copy rule in the network. To optimize the distribution of messages in the network, *ORBOPH* minimizes the message copy amount based on optimal path hops for the omni-directional path/route infection. The following assumptions are made to assure the optimality of *ORBOPH*:

1) The exponential distribution of inter-contact times or at least a distribution with an exponential tail.
2) The independent and identically distributed movement of all nodes in the network.
3) The discovery of the optimal path and infection based on all paths of all nodes.

ORBOPH differs from the existing opportunistic replication

1550-445X/18/$31.00 ©2018 IEEE
DOI 10.1109/AINA.2018.00047

algorithms in the following essential objectives:

1) We model the spreading dynamics of the message based on mobility and paths. Our approach considers the role of the infection states using Markov Chains and network graphs. Our optimal infection is based on the *All Hops Optimal Path (AHOP)* formulation in the system.
2) Our approach provides methodological fundamentals for the study of different Epidemic replication controls and thresholds.
3) The optimality of our model considers undirected paths when it replicates based on the message probability. Also it takes directed paths into account by applying *AHOP* or stopping rules for optimal replication.

Our objective of *ORBOPH* is the following: With some constraints on the optimal replication number for each message, the replication rule of *ORBOPH* increases the efficiency of resource allocation by decreasing the delivery overhead of all messages. Based on the *AHOP* formulation, one concept of *ORBOPH* is to use an optimal stopping rule problem to model each message's replication.

The rest of this paper is organized as follows. In Section II, we first review previous work including the latest developments and relevant results. Then in Section III, we develop our analytical framework by formalizing the actual spreading mechanisms (Section III-A) and model the congestion and dropping policy (Section III-B) for two types of messages (1) by taking into account the buffer size (removed messages) and (2) by considering the message timing (expired messages). Also, we describe our analytical framework mathematically. Based on this framework, we investigate Epidemic's spreading threshold according to the optimal path hops of our OppNet network graph. In Section IV we compare the behavior of different scheduling and dropping policies of Epidemic with our proposed framework. Finally, we conclude our work in Section V.

II. Related Work

In general, the OppNet routing protocols are divided into two main categories, which are the *flooding-based* and the *utility-based* ones. The main goal of utility-based routing is the reduction of resource consumption; resources that are essential in an OppNet are energy, storage and bandwidth. This improvement is achieved by the reduction of replicated messages spread throughout the network. PRoPHET [5] is an utility-based routing protocol which dispatches the messages to neighboring nodes over multiple hops based on delivery predictions. PRoPHET forwards the message via the optimal path based on the calculated probability of delivering the message to the destination. For utility-based routing with limited message copies the *GTMX* [5] forwarding strategy can be used. The routing protocols of the flooding-based category generally copy the message to every encountered node. The price for the redundancy benefit of unlimited message copies is the need for unlimited node resources. One classic flooding-based routing protocol is Epidemic [3]. With Epidemic, messages are replicated to all non-infected (susceptible) nodes whenever encountering one, thus spreading the messages in an omni-directional manner. The mathematical model of Epidemic assumes that the amount of hops on the path to the message destination can be equal to the total amount of nodes in the network. Epidemic is an effective routing protocol when an unlimited buffer is given. However, the buffer size is mostly not sufficient, especially with mobile devices and their physical limitations. This multi-copy approach and the redundancy of paths lead to performance degradation by buffer overflow or congestion. Another well-known DTN routing scheme is Direct Delivery [6] which opposes Epidemic, because it never replicates any messages. Instead, the source node carries the message until it encounters the destination node itself, which results in a hop count of one.

Both, the overhead of a message and the node resource consumption are a problem of Epidemic. Several quota-based routing protocols have been suggested to solve this issue, such as the Spray & Wait [7] and the Spray & Focus [8] scheme. These protocols spread the message with omni-infection fashion, too. However, the total amount of copies that are present in a network is limited to a certain number as they build on both aforementioned protocols (Epidemic and Direct Delivery). The amount of message copies is adapted between two extremal values; a single message as in Direct Delivery and an unlimited number of copies as in Epidemic, which results in a combined benefit of both routing protocols. Both, Spray & Wait and Spray & Focus, limit the amount of message copies to a predefined value L. In Spray & Wait only the source node can *spray* the message copies to L other nodes which then *wait* to deliver the message to the destination themselves. This restricts the maximum hop count of messages to two hops. A modified version of this protocol is Binary Spray & Wait, where each node delivers half of the message copies to encountered nodes. Those that have only one message copy left cannot spread but only deliver the message to the destination, which increases the maximum hop count of messages to $\log L$. Spray & Focus uses its *focus* phase to pass the message copy to a node with a higher probability of delivering the message, here the hop count for messages can go up to N again. The disadvantage of Spray & Wait and Spray & Focus is that the optimal path is determined by the administrator who sets this certain number L of message copies, regardless of the actual network information graph. The analysis and control of routing in DTNs has been studied before. Groenevelt et al. [9] modeled the greedy Epidemic replication and two-hop routing using Markov Chains based on differential equations. The authors inferred a relationship between the message latency and the amount of message copies that are replicated. Furthermore, Zhang et al. [10] implemented a mathematical model using differential equations of homogeneous nodes to analyze Epidemic greedy replication issues. While many researchers in the field apply hop limitations to downsize message replication such as in [11], most of them use a two-hop limit for message spreading control in their routing protocol design [12] [7]. Furthermore, the authors of [13] apply *AHOP* [4] on the DTN environment network graph. This work considers two types of weights to find the optimal path hops using real mobility traces. The authors conclude that the number of hops in an optimal path through static and dynamic network graphs is approximately three or four hops. Neglia and Zhang [14] apply optimal control of flood replication in DTN environments using Epidemic. The paper assumes that the nodes in the entire network know the number of hops that already relayed the message. The optimal control policy is then based on a threshold that is determined by the amount of relay nodes that have a message copy in their buffer storage. The authors in [4] analyze and study the effect of hop count limits on the routing efficiency by modeling the optimal hop limited routing as an *AHOP* problem. The paper considers the bandwidth and delay weights as Quality of Service parameters. In [15], Jia Xu provides an optimal joint expected delay forwarding (OJEDF) scheme. The concept of this scheme is built on the trade off between each single messages' expected delay and the number of message copies for replication. With additional information about the network graph and the optimal path hops, multiple DTN routing schemes try to gain better performance. The Optimal Probabilistic forwarding [16] strategy, for example, tries to increase the delivery rate of the network based on hop count and message life time limitations.

III. System Model and Assumptions

The design of our framework considers a network model that consists of an Ordinary Differential Equation *(ODE)* Markov Chain model, an *All Hops Optimal Path (AHOP)* problem formulation of the network graph, and replication and drop policies for the buffer management. The framework model assumes that at any time, when a new connection between encountered nodes is established, the resource allocation problem arises, where OppNet nodes are suffering from limited resources as bandwidth and storage. This resource allocation problem is inherited by the replication problem where, if a node has messages in its buffer, it must decide if the encountered node is suitable as the next hop, according to the selected routing protocol. Additionally, the node does not know if the connection will last long enough to transfer all messages chosen by the routing protocol. And if this isn't enough, a solution must be found for the problem regarding the order of the messages, because the scheduling of messages is important for increasing the overall delivery probability of the network.

In Table I all the various quantities and notations that are used for the design of our theoretical framework model are summarized. Every node holds the local information of these variables.

TABLE I
Various Quantities and Notations

Variable	Description
N	Amount of network nodes
h_{count}	Number of nodes a message has infected on a path
ICT	Inter-Contact Time
P_{single}	Probability of a message to be replicated
v_{TTL}	Time To Live of message v
$P(t)_\infty$	Probability of a message's delivery
i_t	Number of infected nodes in the network at time t
i_{count}	Number of nodes infected by the current node
i_{opt}	Optimal replication number of a message
h_{opt}	Optimal hop count of a message
$I'(t)$	The rate of infected nodes
P_{opt}	Probability of a message's optimal delivery
P'_{single}	First differentiation of message replication probability
λ_c	Average meeting rate between two nodes in the network
P^{-1}	Inverse of a probability
$D(m)_{max}$	Message to be dropped

The main purpose of our framework *ORBOPH* is the maximization of a network's delivery ratio. Therefor, *ORBOPH* makes its replication decision utilizing local information and optimal path hops of the network graph. The proposed optimal replication control threshold can be calculated by means of a function that evaluates the utility of each message and node in the network. The proposed optimal replication strategy uses this utility value of messages and nodes to decide on the messages' replication. Moreover, the utility of the messages is used to select the message that should be dropped when there is no space left in the node buffer. One of the main issues is the estimation of the ratio between infected and total nodes. With Opportunistic Networks, finding the global number of message copies in the network is a challenge. With local information of each node, which is gained through incoming messages, we can estimate the number of message copies in the network at every time.

The ratio between infected nodes i and total nodes N in the network can be modeled through a finite-state Markov Chain based on the *AHOP* problem. We can approximate this ratio by solving the following Ordinary Differential Equation:

$$I'(t) = \frac{\partial i}{\partial t} = \lambda_c i(N-i) \tag{1}$$

The mobility parameter λ_c specifies how much the nodes move around in the network in terms of the average meeting rate between two nodes. Hence it impacts on the network graph and the messages' states of process. By integrating Eq. (1) we can calculate the number i_t of infected nodes for a point t in time:

$$i_t = \frac{1}{\frac{1}{N} + exp(-tN\lambda_c) - \frac{1}{N}exp(-tN\lambda_c)} \tag{2}$$

We calculate the probability of a message's delivery for the whole system. This probability is expressed as a function over all infected nodes in the entire network as follows:

$$P(t)_\infty \approx \frac{exp(v_{TTL}N\lambda_c) - 1}{exp(v_{TTL}N\lambda_c)} \tag{3}$$

We consider that the meeting rate between nodes λ_c is usually greater than the message copy rate, i.e. a node establishes contacts to other nodes more often than he replicates a message. Therefore, we adapt our system model's function using unidirectional infection and hop count towards the message destination as follows:

$$P_{single} = \frac{\sum_{n=1}^{N} i_n(t)}{N} = \frac{1}{1 + h_{count} + i_{count}} \quad (4)$$

Eq. (4) of our model shows that the probability for a node to replicate the message again, i.e. infect new nodes, depends on two things. First, the message's hop tracking in the network graph and second, the infection counter which resembles the spreading speed of the message copying process. The hop counter h_{count} and the infection or replication counter i_{count} are related to each other by the linear function ($i_{count} = h_{count} + 1$). Alternatively, we can calculate the replication probability of a single message based on Eq. (3) using the mobility parameter λ_c as follows:

$$P_{single} = \frac{exp(\lambda_c v_{TTL}) - 1}{exp(\lambda_c v_{TTL})} \quad (5)$$

Eq. (5) shows that the probability of message replication to a susceptible node depends on the message's life time.

A. Optimal Replication Based on Optimal Path Hops (ORBOPH)

OppNets are specified as a network of mobile nodes which can exchange messages during their connection period. Nodes can connect when they are in each others coverage area. The concept of a hop-limited routing scheme is that each message has a hop count field, which shows the maximum number of infected nodes or hops this message can pass through. When a message is replicated or transmitted from one node to another, the hop count field of this message is incremented by 1 at the receiving node. If the hop limitation of the message is h, the message can only be replicated or transmitted to h hops. Also, in OppNets, the topology graph changes frequently, as a result of the mobility model. However, node contacts that are sustained for a long time can be used to estimate future contact probability and hence the network topology graph. For this network graph, given a source node has a message which should be sent to a destination via a specific path, our model aims to find the shortest path or rather the minimal number of message transmissions to the destination. When a path has at most k hops, for each k, $1 < k < h$, this problem can be formulated in terms of *AHOP*.

In this section, we derive the probability for the optimal delivery P_{opt} of a single message. We state that this optimal delivery is achieved when the message's hop count h_{count} is equal to the optimal path hops value h_{opt}, i.e. when the message passed h_{opt} nodes on it's path to the current node. This particularly applies to the destination node. Therefore, the optimal path hops value h_{opt} is the upper bound of the message's hop count h_{count}. Furthermore, we consider the fact that Epidemic is based on an omni-directional replication process, whereas our model aims to perform the replication in a directional manner as routed path based on the *AHOP* formulation.

Our assumption is based on the mathematical circle area $(i_{opt}) = {h_{opt}}^2 \cdot \pi$, i.e. we use h_{opt} as radius, as the message is spread in all directions. We propose that for achieving the maximum delivery using a single path instead of omni-directional replication the optimal infection or optimal replication i_{opt} can be assumed as:

$$i_{opt} = (h_{opt} + 1)^2 \quad (6)$$

The further a message has to travel, the less likely is an optimal delivery. Based on Eq. (4) and Eq. (5) we now determine the probability for optimal delivery P_{opt} of each message utilizing the optimal hop count and the optimal infection:

$$P_{opt} = \frac{1}{1 + h_{opt} + i_{opt}} \quad (7)$$

Each time an infected node finds a susceptible node in his direct neighborhood, it has to respect some criteria to decide whether to forward the message to this susceptible node or not. The optimization objective is to minimize the costs, which is achieved by the adherence to the optimal path hops. Consequently, we want to characterize the Markov decision problem that targets to find the probability for optimal delivery and can be computed with the following equation:

$$P_{opt} = P'_{single} \quad (8)$$

The right side of Eq. (8) can be substituted for the differentiation of Eq. (5) as follows:

$$P_{opt} = \lambda_c v_{TTL} exp(-\lambda_c v_{TTL}) \quad (9)$$

As the mobility parameter λ_c goes to zero, the inter-contact time that passes between two meetings of a pair of nodes increases. This leads to a network graph that is more sparse rather than connected. Then we can proof that:

$$\lim_{\lambda_c \to 0} exp(-\lambda_c v_{TTL}) \approx 1 \quad (10)$$

Now from Eq. (6), Eq. (7) and Eq. (9) the optimality is based on the mobility parameter λ_c and the network graph. Furthermore, the probability of optimal delivery considers i_{count} in the message replication process. In addition, it considers the hop count h_{count} as a path direction of the network graph. We can now describe the probability of delivery optimality with the following function:

$$P_{opt} = \lambda_c v_{TTL} = \frac{1}{1 + h_{opt} + (h_{opt} + 1)^2} \quad (11)$$

Our model considers $\lambda_c v_{TTL}$ to be constant during the same scenario. Factoring out Eq. (11) we get the following function:

$$\lambda_c v_{TTL} = \frac{1}{2 + 3h_{opt} + {h_{opt}}^2} \quad (12)$$

Now from Eq. (12) we find that the optimal path hops value h_{opt} is based on the probability of optimal delivery value that is $\lambda_c v_{TTL}$. In addition, for our probability we consider the global network graph path h_{count}. Also, the probability of optimal delivery considers the local message infection per node i_{count} as follows:

$$\lambda_c v_{TTL} = \frac{1}{(h_{opt}+1)(h_{opt}+2)} \tag{13}$$

The previous Eq. (13) expresses that, for starters, the most favorable policy is the infection of every susceptible node that is in range. But the infection should stop when the amount of infected nodes has reached $(i_{opt})^2+(i_{opt})^{0.5}$. Another criterion for the stoppage of infection is the successful delivery of the message to the destination. Finally, with Eq. (6) and Eq. (13), we rewrite the relationship between optimal path hops value h_{opt} and the optimal infection value. This relationship shows that the optimal infection is a square of optimal hops as follows:

$$\frac{1}{\lambda_c v_{TTL}} = 2 + 3h_{opt} + h_{opt}^2 = i_{opt} + i_{opt}^{0.5} \tag{14}$$

Eq. (14) states that the inverse of the probability of optimal delivery for our model is considered as a second order function in h_{opt}. The equation also states that it is considered as first order function in i_{opt}. This shows that i_{opt} can be calculated based on a square of h_{opt}.

B. Buffer Occupancy Improvement

For the improvement of the buffer control we presume that all nodes share the same traffic parameters and infection strategy. In this section we will implement our drop policy approach which decides which message shall be evicted from the buffer when it is full. In order to do this, we will implement decision utilities for infection and dropping based on their optimality. As the main goal of our model is to control the replication of messages and the infection of nodes, our model will implement this utility for infection. It uses the send queue policy based on single message probability. Furthermore, the control policy determines the threshold decision of message copies and infected nodes. The calculation of this threshold is based on the probabilities of replication and optimal delivery as follows:

$$send-queue = \begin{cases} MaxP_{single} & P_{single} > P_{opt} \\ 0 & otherwise \end{cases} \tag{15}$$

For performance improvements of our system model, there are two types of buffer evicted messages. The first occasion in which a message is dropped occurs when the buffer is full. The decision which message will be deleted is made according to the following formula:

$$D(m)_{max} = P_{single}^{-1} \tag{16}$$

The message least likely to be replicated is dropped. As Eq. (16) states, by using the inverse of the probability of a single message being replicated, the drop decision depends on the optimal replication i_{opt} of the current message carrier. The drop policy also takes into account the meeting rate λ_c of encountered nodes which are impacted by the mobility as well as the number of copies created to cross the path of length h_{opt}. The optimal hops are calculated as shorted path between the source and the final destination that is found based on the network graph. The three values i_{opt}, λ_c and h_{opt} are regarded as main function parameters of the probability of optimal delivery. They can also be used as variables of the threshold function for the maximum number of message copies.

$$Expired = \begin{cases} P_{single} <= P_{opt} & i_{count} > i_{opt} \\ P_{single} <= P_{opt} & h_{count} > h_{opt} \end{cases} \tag{17}$$

The second occasion in which a message is dropped occurs when it expires. Eq. (17) shows that the expiration of a message can be triggered by two cases. On one hand the number of replicates of the message (i_{count}) can already be higher than the optimal number of replicates (i_{opt}). On the other hand the message might have passed more hops (h_{count}) than would be optimal (h_{opt}). Finally, apart from Eq. (17), whether a message is expired can also be determined by the Time-To-Live (TTL) of a message. The node periodically checks the TTL value of each message. If any message has a value which is equal or less than zero it is declared expired.

IV. Experiment and Results

To evaluate the proposed framework *ORBOPH*, it is required to determine the metrics. These metrics will be used for the comparison of *ORBOPH* with other references. The performance of the proposed *ORBOPH* will be compared with different drop and replication policies, and different buffer managements. We set up the Epidemic router as a greedy infection scheme with different replication and drop polices. The proposed framework *ORBOPH* is evaluated based on different types of messages and network graph information. We configure drop approaches using the number of nodes i_{count} infected by a certain node as spreading speed of the message copy process. Also our evaluation considers number of infected nodes h_{count} on a message's path as indication of this path. Furthermore, h_{count} shows the path of unidirectional links to the destination. The following metrics are used in the comparison and evaluation:

1) *Delivery Ratio* is the number of destination nodes which successfully receive a message in relation to the total number of infected and source nodes of the messages.
2) *Overhead Ratio* is the amount of infected nodes used for transferring one message to it's final destination in relation to the total number of nodes.
3) *Average Delay* is the average time that the successfully delivered messages spent in the buffers of infected nodes on their way.

A. Data and Experimental Settings

The performance of *ORBOPH* is evaluated using the ONE Simulator [17], [18]. The parameters and settings for our evaluation scenarios are listed in Table II. This comparison considers the performance of our model in view of the three metrics that have been presented in the last section. As a reference, the routing protocol Epidemic was used with two replication strategies for the send queue and three drop policies for the buffer management.

The first replication strategy is based on the number of nodes infected by the considered node (i_{min}) and the second strategy is based on the number of infected nodes on the path through the network (h_{min}).

For *ORBOPH*, we use the replication probability of the message for the replication stopping rule as explained in Eq. (14). From the replication stopping rule condition we derive the value 20 as the maximum number of infected nodes. Now we can draft the probability of optimal delivery based on the message information and the network graph by using Eq. (14).

For the priority of the send queue (Eq. (15)) we use a sorting based on i_{count} and h_{count} i.e. based on the number of nodes infected by the current node and the number of nodes the message passed on it's path. This is because every message will be transmitted as a replication (i_{count}) but also stored in the buffer (h_{count}).

For the buffer management and drop polices we consider the following techniques for comparison (Eq. (16) and Eq. (17)). First and second, the maximum number of replications (i_{max}) and hops (h_{max}). Third, MOFO (Evict Most Forwarded First), which considers the number of replications as the main decision criteria and the number of hops as a tie breaker.

For our *ORBOPH* we apply a different drop policy as shown in Eq. (16). We consider the optimal number of infections and path hops to improve the performance of *ORBOPH*.

For a first result, that will be necessary for the following scenarios, we ran the scenarios for measurements of the inter-contact time (ICT) between all encountered homogeneous nodes and found an ICT of approximately 5988 seconds. From the measured ICT, we calculate the inter-meeting rate of the node pairs in the system. This value, λ_c, was 0.000167. In our experiment scenarios h_{opt} is equal to 3 hops. Furthermore, we calculate the value of optimal replication $i_{opt} = 16$ from Eq. (6). From Eq. (11), we now can calculate the probability of the optimal delivery of each message. This probability is equal to 0.0501 based on the optimal path hops i_{opt}.

These values are used as described in Eq. (15), Eq. (16) and Eq. (17) for the implementation and therefore for all scenarios in the following evaluation.

B. Numerical Results

During our simulation scenarios, we vary the message TTL between 100 to 500 minutes to simulate different traffic loads and buffer congestion states. We show results of three metrics which are delivery ratio, overhead and delay over a variety of infection and drop policies, and buffer management options. For comparison, we use the following three different scenarios:

For the first scenario we select the number of infected nodes as one of the main decision parameter of the buffer management, where the message which has the minimal replication amount i_{min} will be replicated first. Regarding the drop decision the message with the highest infection rate i_{max} will be dropped first.

TABLE II
SIMULATION SETTINGS

Simulation area	Helsinki, Finland Map
Simulation time	12h
Number of devices (N)	126
Group type / speed	80 Pedestrians (0.5 to 1.5km/h) 40 Cars, 6 Trains (10 to 80km/h)
Routing protocols	Epidemic
Interface type	Simple Broadcast
Transmission range	250m
Bandwidth	250KBps
Replication strategies used	Min Replication (i_{min}) Min Hops (h_{min})
Drop policies used	Max Replication (i_{max}) Max Hops (h_{max}) MOFO
Message sizes ranges	500KB, 1MB
Message creation interval	25, 35s
Time-To-Live (TTL)	100, 200, 300, 400, 500min
Default buffer size	5MB (Pedestrians) 50MB (Cars, Trains)

As another main decision of the buffer management we select the number of path hops, where the message which has the shortest path h_{min} will be replicated first. Regarding the drop decision the message with the longest path h_{max} will be dropped first.

For the second scenario we consider i_{min} and h_{min} as scheduling decisions, but apply MOFO as the drop policy which considers i_{max} as the main decision and h_{max} as a tie breaker.

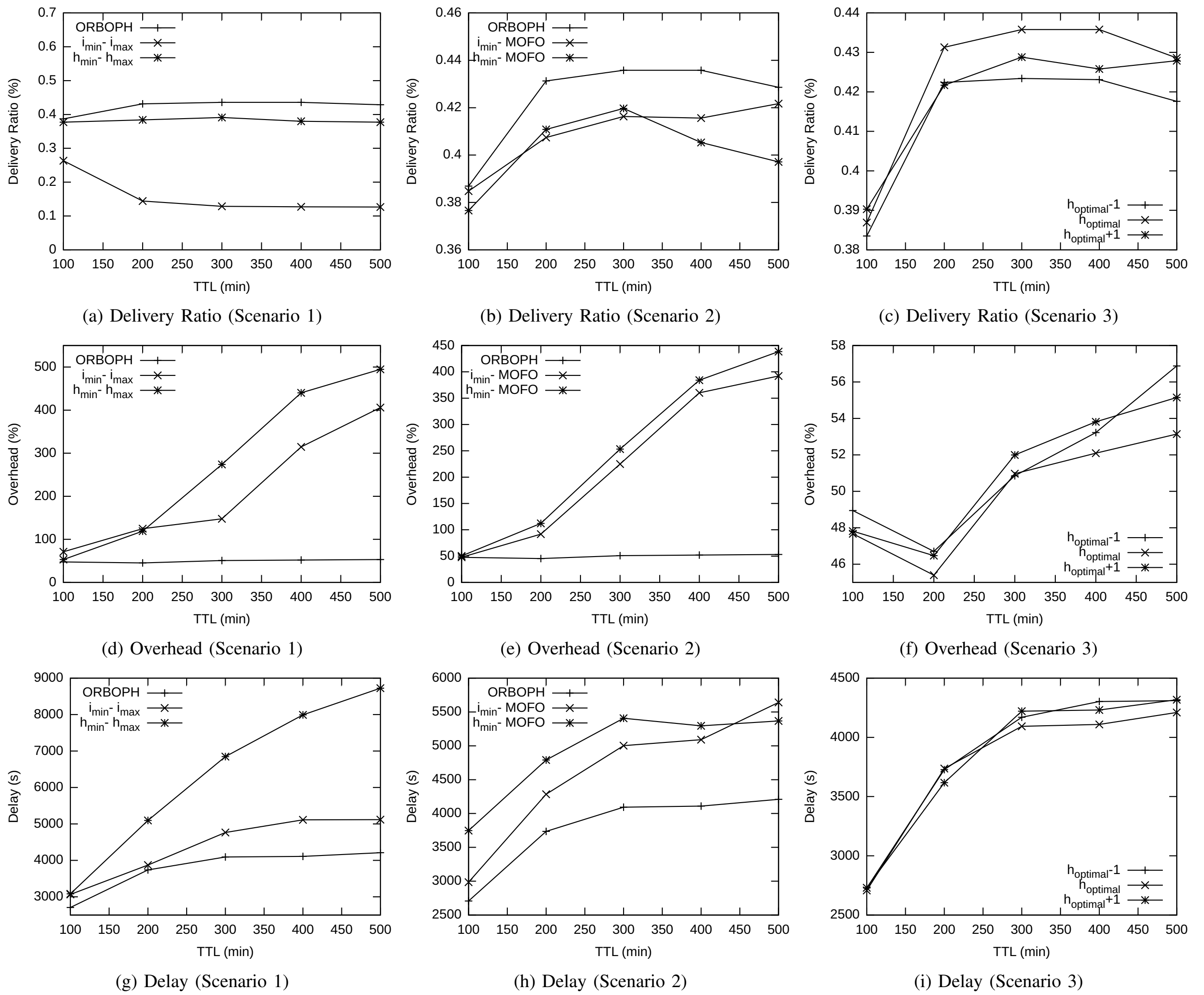
For the third scenario we investigate the impact of the h_{opt} value on our proposed model *ORBOPH*. We consider only the h_{opt} value, because the i_{opt} value is a function of h_{opt} as shown in Eq. (6).

Finally, we compare our proposed *ORBOPH* with other models within those three different scenarios.

1) Delivery Ratio: In the first scenario the delivery ratio of *ORBOPH* is compared with two options that use message scheduling based on the minimum i_{count} and h_{count} for their send queue and with different dropping policies which are based on maximum i_{count} and h_{count}. Furthermore, in the second scenario we consider the MOFO policy which combines the two criteria in priority of infection and then the path hops. The comparisons are depicted in Figure 1(a), (b) and (c).

From Figure 1(a) it can be seen that the delivery ratio of *ORBOPH* is higher than those of Epidemic with the applied message scheduling. The improvement of the delivery ratio reaches 8% compared with the h_{count} decision and more than 30% compared with the i_{count} decision at a TTL of 500min. This is because the optimality of the message delivery is a function of both optimal path hops and infection.

In the performance evaluation of *ORBOPH* for the second scenario, in which MOFO is applied, the delivery ratio of *ORBOPH* is still higher (($\approx 1 - 4\%$)) than compared with the others (Figure 1(c)). The similar values of the compared



(a) Delivery Ratio (Scenario 1) (b) Delivery Ratio (Scenario 2) (c) Delivery Ratio (Scenario 3)

(d) Overhead (Scenario 1) (e) Overhead (Scenario 2) (f) Overhead (Scenario 3)

(g) Delay (Scenario 1) (h) Delay (Scenario 2) (i) Delay (Scenario 3)

Fig. 1. Measurement results with regards to the three scenarios

approaches derives from applying MOFO, which considers both i_{count} and h_{count}.

For the third scenario, shown in Figure 1(c), we compare the delivery ratio of *ORBOPH* with the optimal path hops value increased or decreased by one. We can see that the optimal path hops value, which is, according to Eq. (14), equal to 3, can be confirmed.

From Figure 1(a) we find that Epidemic, based on an infected nodes decision i_{count}, reaches a low delivery ratio. We believe this to happen because the optimal path hops are a main component of the probability of optimal message delivery. Therefore, to improve our model, we consider the infection value i_{count} as a function of optimal path hops h_{opt} as shown in Eq. (6).

2) Overhead Ratio: Overhead Ratio is one of the main performance factors to compare *ORBOPH* with other buffer management policies of Epidemic because this factor is related to the infection value i_{count} and the path hops value h_{count}.

Therefore, we consider this factor as our main factor, because the main idea of *ORBOPH* is to consider the resource consumption in terms of storage, path hops (h_{count}) and transmission (i_{count}).

Figures 1(d) and (e) show that *ORBOPH* has lower overhead ($(\approx 40\%)$) when compared with the other scenarios. This is because the send queue policy for *ORBOPH* is based on an optimal message delivery probability. This probability is restrained by the replication stopping rule which on the other hand is based on the optimal path hops of the network graph.

This optimal message delivery probability is calculated as a function of the number of path hops h_{count} as a storage metric and the replication counter i_{count} as a transmission metric to stop the replication of a single message. This stopping rule also depends on the optimal path hops h_{opt}.

Furthermore, we notice that *ORBOPH* has a more stable overhead ratio, because the probability of the stopping rule for infection has the same condition as the applied dropping policy. To accomplish the desired message delivery, the main concept of infection and dropping choices should be considered as a trade-off. In general, while the node infection and message spreading will indeed improve the delivery ratio, we should note that the dropping rate of the message will increase with lower availability of limited node buffer. This limitation of storage in the node buffers leads to high overhead and resource consumption.

Furthermore, in Scenario 3, depicted in Figure 1(f), we find that *ORBOPH*, with the applied optimal path hops formulation, still has the lowest overhead ratio when compared to the increased or decreased optimal path hops.

3) Delay: The delay metric, here as an average end-to-end delay from source to destination, is commonly considered as a performance metric for applications and scenarios in DTNs and OppNets.

In Scenario 1, depicted in Figure 1(g), we examine that *ORBOPH* has a lower delay when it considers the optimal path hops as infection and dropping decisions, especially when there are many messages in the buffer as Eq. (16) and Eq. (17) show. We believe that the number of path hops h_{count} is the main factor of both, this delay and the overhead from before. This is because the optimal infection i_{opt} is a function of h_{opt} and the *AHOP* formulation used by our framework as a buffer management decision impacts on the average end-to-end delay. We can see that by changing the infection and dropping criteria as shown in Figure 1(h) in Scenario 2, *ORBOPH* has a lower delay ($\approx$ 1000s less at a TTL of 500min) when compared with the the reference implementations.

Furthermore, Figure 1(i) of the third scenario shows that *ORBOPH* also has a low delay when using the optimal path hops instead of increased or decreased ones.

V. Conclusion and Future Work

This paper aims to formulate the problems of Epidemic replication control as a Markov Chain model with consideration of the network graph based on the *All Hops Optimal Path (AHOP)* problem. The paper considers the infection of Epidemic as a function of optimal path hops for obtaining the probability of optimal message delivery under different buffer managements. We solve the problem by combining an *Ordinary Differential Equation (ODE)* approximation and *All Hops Optimal Path (AHOP)*. The proposed *ORBOPH* is constructed based on analytical studies which consider the optimality of omni-infection as a function of optimal path hops. This means that the basic omni-infection of Epidemic can be controlled by the directional message transgression towards the destination. The numerical results of this paper show that optimal node infection or message replication of Epidemic is the square of the optimal path hops to achieve better performance. For future work, we suggest to extend the research to consider the timing of messages such as buffer time for minimizing Epidemic delay. Another option is to introduce a distributed monitoring approach [19] to learn about the network and to adapt to it, or to consider social relations through short term and long term interactions such as in [20].

References

[1] A. Ippisch and K. Graffi, "Infrastructure mode based opportunistic networks on android devices," in *2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA)*, 2017, pp. 454–461.

[2] I. Woungang, S. K. Dhurandher, A. Anpalagan, and A. V. Vasilakos, *Routing in Opportunistic Networks*. Springer Publishing Company, Incorporated, 2013.

[3] A. Vahdat and D. Becker, "Epidemic Routing for Partially-Connected Ad Hoc Networks," *Technique Report, Department of Computer Science, Duke University, USA*, vol. vol.20, no. 6, 2000.

[4] R. Guérin and A. Orda, "Computing shortest paths for any number of hops," *IEEE/ACM Trans. Netw.*, vol. 10, no. 5, pp. 613–620, 2002.

[5] A. Lindgren, A. Doria, E. Davies, and S. Grasic, "RFC 6693: Probabilistic Routing Protocol for Intermittently Connected Networks," *IETF*, 2012.

[6] A. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Single-Copy Routing in Intermittently Connected Mobile Networks," in *Proceedings of the Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON)*, 2004, pp. 235–244.

[7] T. Spyropoulos and K. Psounis, "Spray and Wait: An Efficient Routing Scheme for Intermittently Connected Mobile Networks," *Proceedings of the ACM SIGCOMM Workshop on Delay-Tolerant Networking*, 2005.

[8] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and Focus: Efficient Mobility-Assisted Routing for Heterogeneous and Correlated Mobility," in *In Proceedings of the IEEE International Conference on Pervasive Computing and Communications - Workshops (PerCom Workshops)*, 2007, pp. 79–85.

[9] R. Groenevelt, P. Nain, and G. Koole, "The Message Delay in Mobile Ad Hoc Networks," *Performance Evaluation*, vol. 62, no. 1-4, 2005.

[10] X. Zhang, G. Neglia, J. F. Kurose, and D. F. Towsley, "Performance Modeling of Epidemic Routing," *Computer Networks*, vol. 51, no. 10, pp. 2867–2891, 2007.

[11] M. Vojnovic and A. Proutière, "Hop limited flooding over dynamic networks," in *INFOCOM 2011. 30th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, 10-15 April 2011, Shanghai, China*, 2011, pp. 685–693.

[12] A. A. Hanbali, P. Nain, and E. Altman, "Performance of ad hoc networks with two-hop relay routing and limited packet lifetime (extended version)," *Perform. Eval.*, vol. 65, no. 6-7, pp. 463–483, 2008.

[13] S. Bayhan, E. Hyytiä, J. Kangasharju, and J. Ott, "Analysis of hop limit in opportunistic networks by static and time-aggregated graphs," in *2015 IEEE International Conference on Communications, ICC 2015, London, United Kingdom, June 8-12, 2015*, 2015, pp. 3287–3292.

[14] G. Neglia and X. Zhang, "Optimal Delay-power Tradeoff in Sparse Delay Tolerant Networks: A Preliminary Study," in *Proceedings of the SIGCOMM Workshop on Challenged Networks, (CHANTS@SIGCOMM)*, 2006, pp. 237–244.

[15] C. Liu and J.Wu, " An Optimal Probabilistic Forwarding Protocol in Delay Tolerant Networks," *International Journal of Distributed Sensor Networks* , pp. 105–114, 2009.

[16] J. Xu, X. Feng, W. Yang, R. Wang, and B. Q. Han, "Optimal Joint Expected Delay Forwarding in Delay Tolerant Networks," *International Journal of Distributed Sensor Networks*, vol. 9, no. 11, p. 941473, 2013.

[17] A. Keränen, J. Ott, and T. Kärkkäinen, "The ONE Simulator for DTN Protocol Evaluation," in *Proceedings of the 2Nd International Conference on Simulation Tools and Techniques*, ser. Simutools '09, 2009, pp. 55:1–55:10.

[18] A. Keränen, T. Kärkkäinen, and J. Ott, "Simulating Mobility and DTNs with the ONE (Invited Paper)," *Journal of Communications*, vol. 5, no. 2, pp. 92–105, 2010.

[19] K. Graffi, "Monitoring and Management of Peer-to-Peer Systems," Ph.D. dissertation, Technische Universität, Darmstadt, August 2010. [Online]. Available: http://tuprints.ulb.tu-darmstadt.de/2248/

[20] B. Guidi, T. Amft, A. D. Salve, K. Graffi, and L. Ricci, "DiDuSoNet: A P2P Architecture for Distributed Dunbar-based Social Networks," *Peer-to-Peer Networking and Applications*, vol. 9, no. 6, pp. 1177–1194, 2016.